# DAIT/Rho STATISTICAL ANALYSIS PLAN
# 09 November 2020

**ATOPIC DERMATITIS RESEARCH NETWORK**

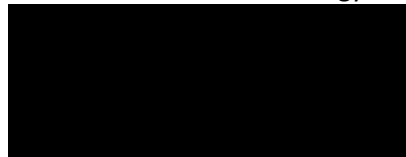**PROTOCOL-ADRN 09**

**Effect of Dupilumab (anti-IL4Rα) on the Host-Microbe Interface in Atopic Dermatitis**

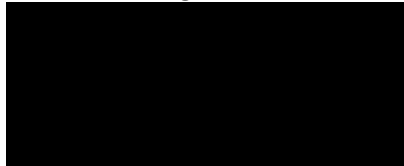**Dupilumab Study**

**VERSION:**       **1.0**

**DATE:**       **09 NOV 2020**

**SPONSOR:**      Division of Allergy, Immunology, and Transplantation – DAIT
National Institute of Allergy and Infectious Diseases – NIH

**PREPARED BY:**    Rho
Rho Building

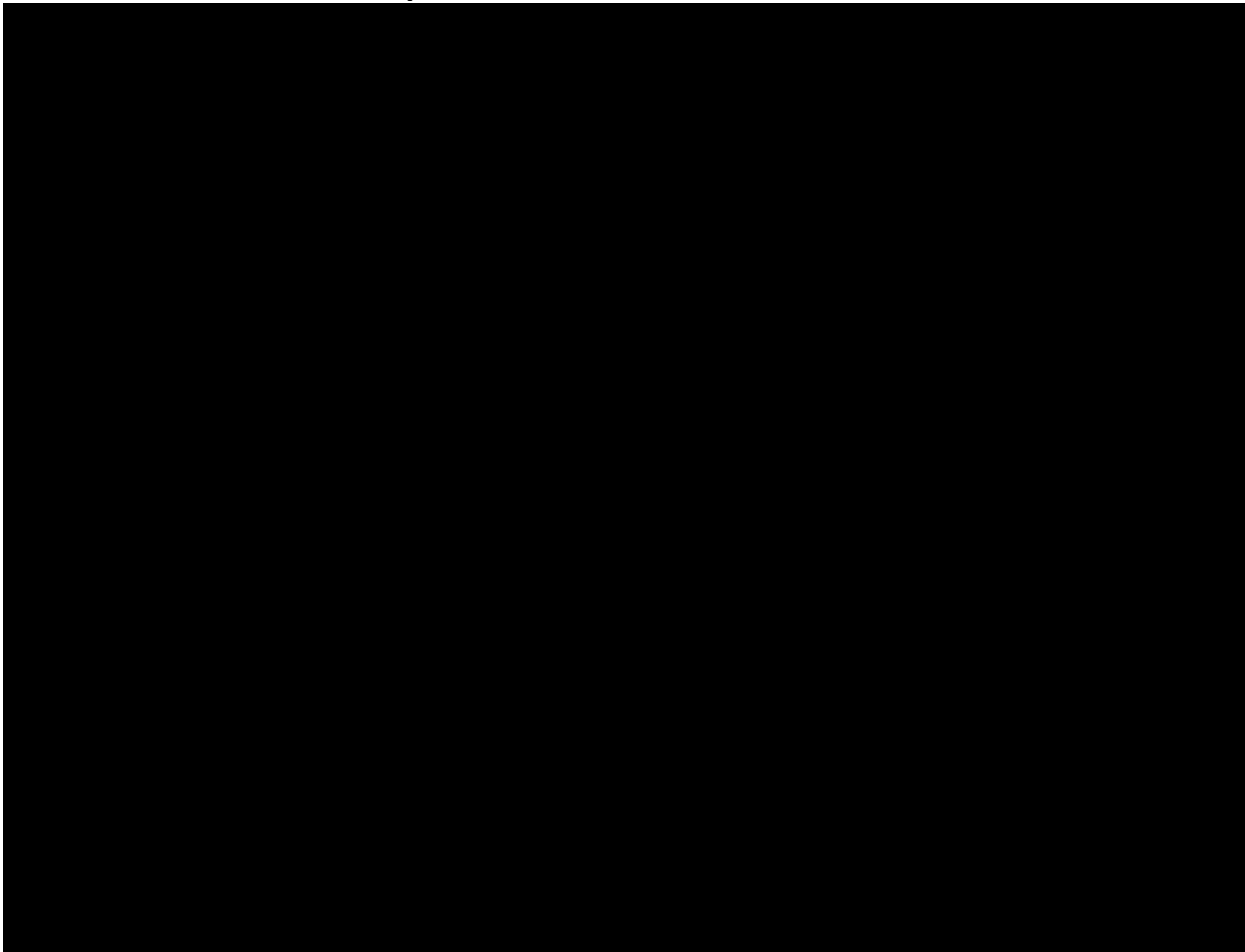# DAIT/Rho STATISTICAL ANALYSIS PLAN ACKNOWLEDGMENT AND SIGNATURE SHEET

**Dupilumab**

Approved:

Approved:

Approved:

Approved:

## Document History

| Version | Date | Comments/Changes |
|---------|------|------------------|
| 0.1 | 25 AUG 2020 | Initial Draft Version |
| 0.2 | 16 OCT 2020 | Revised Draft Version |
| 0.3 | 04 NOV 2020 | Revised Draft Version |
| 1.0 | 09 NOV 2020 | Initial Version |

Table of Contents

## GLOSSARY OF ABBREVIATIONS

| | |
|---|---|
| AD | atopic dermatitis |
| ADA | anti-drug antibodies |
| AE | adverse event |
| AUC | area under the curve |
| BMI | body mass index |
| CBC | complete blood count |
| CoNS | coagulase-negative Staphylococci |
| COVID-19 | Coronavirus Disease 2019 |
| CRF | case report form |
| CSR | clinical study report |
| DAIT | Division of Allergy, Immunology, and Transplantation |
| DNA | deoxyribonucleic acid |
| DSMB | Data Safety Monitoring Board |
| EASI | Eczema Area and Severity Index |
| ECG | electrocardiogram |
| eCRF | electronic case report form |
| ICH | International Conference on Harmonisation |
| ID | participant identifier |
| IGA | Investigator Global Assessment |
| IRB | Institutional Review Board |
| ITT | intent-to-treat |
| max | maximum |
| MedDRA | Medical Dictionary for Regulatory Activities |
| min | minimum |
| mITT | modified intent-to-treat |
| n | sample size |
| NCI-CTCAE | National Cancer Institute-Common Terminology Criteria for Adverse Events |
| NIAID | National Institute of Allergy and Infectious Diseases |
| NIH | National Institutes of Health |
| NRS | numerical rating scale |
| PBMC | peripheral blood mononuclear cell |
| PP | per-protocol |
| RDBPC | randomized double-blind placebo-controlled |

| S. aureus | Staphylococcus aureus |
|-----------|----------------------|
| SACCC | Statistical And Clinical Coordinating Center |
| SAE | Serious Adverse Event |
| SAg | superantigen |
| SAP | Statistical Analysis Plan |
| SAR | Suspected Adverse Reaction |
| SCORAD | SCORing Atopic Dermatitis |
| SD | standard deviation |
| SNP | single nucleotide polymorphism |
| SOC | system organ class |
| TARC | thymus and activation-regulated chemokine |
| TEAE | treatment-emergent AE |
| TEWL | transepidermal water loss |
| Th2 | T helper 2 |
| TJs | tight junctions |
| WHO | World Health Organization |

## 1.    PROTOCOL SYNOPSIS

The protocol synopsis is available in the most recent Institutional Review Board (IRB)-approved version of the Dupilumab protocol (Version 3.0; 29 August 2019).

## 2.    INTRODUCTION

This Statistical Analysis Plan (SAP) is based on Section 13 *Statistical Considerations and Analytical Plan* of the Dupilumab protocol. However, the analyses specified in this document supersede the high-level analysis plan described in the protocol. The SAP contains detailed information to aid in the implementation of the statistical analyses and reporting of the study results in the clinical study report (CSR). This SAP was written with consideration of the recommendations outlined in the most recent International Conference on Harmonisation (ICH) E9 Guideline, entitled Guidance for Industry: Statistical Principles for Clinical Trials and the most recent ICH E3 Guideline, entitled Guidance for Industry: Structure and Content of Clinical Study Reports.

## 3.    GENERAL ANALYSIS AND REPORTING CONVENTIONS

The following analysis and reporting conventions will be used:
- Categorical variables will be summarized using counts (n) and percentages (%) and will be presented in the form "n (%)."  Percentages will be rounded to one decimal place. If a count is 0, 0% will be shown for the percentage.
- Continuous variables will be summarized using sample size (n), mean, standard deviation (SD), median, first and third quartiles, minimum (min), and maximum (max), as appropriate.  The mean will be reported at one more significant digit than the precision of the data, the standard deviation will be reported at two more significant digits than the precision of the data, and first and third quartiles, minimum, maximum, and median will be reported at the same precision as the data. The median will be reported as the average of the two middle numbers if the dataset contains an even number of observations.
- Test statistics including t, z, and χ2 test statistics will be reported to two decimal places.
- P-values will be reported to three decimal places if greater than or equal to 0.001.  If less than 0.001, the value will be reported as "<0.001."  A p-value can be reported as "1.000" only if it is exactly 1.000 without rounding.  A p-value can be reported as "0.000" only if it is exactly 0.000 without rounding.
- Each data listing will be sorted by treatment arm and participant identifier (ID) unless otherwise noted.
- All analyses described in the SAP will be performed using SAS® System Version 9.4 or later or R Version 3.5 or later.

If departures from these general conventions are presented in the specific evaluations section of this SAP, then those conventions will take precedence over these general conventions.

## 4.    ANALYSIS POPULATIONS

**Intent-to-treat (ITT) population**: All participants who were randomized. Participants will be analyzed according to the treatment arm to which they were randomized, regardless of the medication they actually received.

**Modified intent-to-treat (mITT) population**: All participants who were randomized and have *Staphylococcus aureus* (*S. aureus)* abundance measured at Day 0 and Day 28. Participants will be analyzed according to the treatment arm to which they were randomized, regardless of the medication they actually received.

**Safety population**: All participants who were randomized and received at least one dose of study treatment. Participants will be analyzed according to the medication they actually received, defined as "dupilumab" if a participant received any dose of dupilumab and "placebo for dupilumab" otherwise, regardless of the treatment arm to which they were randomized. Non-treatment emergent adverse events (e.g., any adverse event (AE) that occurs before the first injection) will be summarized in all participants who are enrolled, while treatment emergent adverse events (TEAE) (e.g., any AE that occurs on or after the first injection) will be summarized in the safety population.

**Per-protocol (PP) population**: All participants who were randomized, received the correct medication and dosage to which they were randomized, did not have a major protocol deviation, did not miss any scheduled injections, did not used any of the medications/procedures listed in Protocol Section 7.3 with the exception of topical calcineurin inhibitors (tacrolimus or pimecrolimus), topical corticosteroids, topical phosphodiesterase inhibitors (crisaborole) and topical antibiotics if applied to areas of the skin not being sampled, and have met sampling criteria.

## 5.    STUDY PARTICIPANTS

### 5.1.    Disposition of Participants

The disposition of all enrolled participants will be summarized in tables and listed. The following disposition information will be summarized:

- The number of participants screened.
- Among screened participants:
    - The number and percentage of participants who failed Screening, including those who attended the Screening Visit but did not meet the inclusion/exclusion criteria.
- The number of participants randomized.
- Among randomized participants:
    - The number and percentage of participants in the ITT population.
    - The number and percentage of participants in the mITT population.
    - The number and percentage of participants in the safety population.
    - The number and percentage of participants in the PP population.
- Among participants in the mITT population and grouped by treatment arm:

- o   The number and percentage of participants in the safety population.
- o   The number and percentage of participants in the PP population.
- o   The number and percentage of participants who completed Day 42 (random double-blind placebo-controlled, RDBPC).
- o   The number and percentage of participants who completed Day 112 (Open-label).
- o   The number and percentage of participants who completed Day 182 (Follow-up).
- o   The number and percentage of participants who withdrew during the study due to:
  - ▪   Withdrawal by participant
  - ▪   Lost to follow-up
  - ▪   Death
  - ▪   Pregnancy
  - ▪   Physician decision
  - ▪   Participant experienced an anaphylactic reaction to study drug injection
  - ▪   Participant was diagnosed with an exclusionary malignancy
  - ▪   Participant acquired an opportunistic infection
  - ▪   Non-compliance with the study protocol
  - ▪   Participant refused protocol-mandated procedure at the Screening or Treatment Initiation Visit
  - ▪   AE
  - ▪   Coronavirus Disease 2019 (COVID-19)
  - ▪   Other
- o   The number and percentage of participants who discontinued study treatment.
- o   The number and percentage of participants who missed visits due to COVID-19.
- o   The number and percentage of participants who had remote visits due to COVID-19.

A data listing of participant disposition will be provided for the ITT population and will include an indication of whether the participant is in the mITT population.

## 5.2.    Demographic and Other Baseline Characteristics

Descriptive statistics for demographic and baseline characteristics will be summarized for the mITT population grouped by randomized treatment arm and overall.

Demographic data will include age at randomization, sex, race, ethnicity, and clinical site.

Baseline will be defined as the last measurement prior to or on the day of randomization (beginning of Day 0). Baseline characteristics will include:

- •   Body weight
- •   Height
- •   Body Mass Index (BMI) percentile
- •   Atopic Dermatitis (AD) Severity
  - o   SCORing Atopic Dermatitis (SCORAD)
  - o   Eczema Area and Severity Index (EASI) Score

- o EASI Score < 21.1
- o Nottingham Eczema Severity Score
- o Investigator Global Assessment (IGA)
- Average itch in the past 24 hours as assessed by Pruritus Numerical Rating Scale (NRS)
- Self-reported history of Eczema Herpeticum
- Self-reported history of Staph infection
- Self-reported history of food allergy
- Self-reported history of animal allergies
- Self-reported pets living in the participant's home
- Self-reported smoking in the participant's home

Demographic and baseline characteristics for continuous variables will be summarized using the number of participants, mean, SD, median, and first and third quartiles, and min and max (as applicable). Demographic and baseline characteristics for qualitative or categorical variables will be summarized by the number and percentage of participants within each category.

A data listing of demographics, baseline characteristics, and medical history will be provided for the mITT population. A data listing of any inclusion/exclusion criteria not met will also be provided for all screened participants, as well as for all participants in the mITT population separately. A data listing of all participants who had visits that were remote or missed due to COVID-19 will be provided for the mITT population and will include the number of visits that were remote or missed due to COVID-19 for each participant.

## 6. STUDY OPERATIONS

### 6.1. Protocol Deviations

Protocol deviations will be recorded on the protocol deviation electronic case report form (eCRF) page. Major protocol deviations will be identified by the DAIT Medical Monitor and study team prior to unblinding of treatment assignment and database freeze/lock. A data listing of all protocol deviations sorted by site and participant ID will be provided to the DAIT Medical Monitor and study team for the manual identification of major protocol deviations.

Following database freeze/lock, a data listing will be provided for all protocol deviations. The listing will be sorted by site, treatment arm, and participant ID and will also include type of deviation, severity of the deviation (major or non-major), date of occurrence, and the reason for the deviation.

Protocol deviations will be summarized in tabular format by treatment arm and by type of deviation (including major deviations).

### 6.2. Treatment Adherence

The number and percentage of participants receiving each dose and any dose of blinded study drug will be summarized by treatment arm in the safety population. The overall duration of exposure

(weeks) and the percentage of expected doses received will be summarized by descriptive statistics (mean, SD, median, first and third quartiles, min and max). Duration of exposure will be defined based on the difference (in days) between the dates of the first and last dose of blinded study drug plus 1 day and then converted to weeks. There are 3 expected doses during the RDBPC portion of the study.

Similarly, the number and percentage of participants receiving each dose and any dose of open-label study drug will be summarized in the safety population. Duration of exposure will be defined based on the difference (in days) between the dates of the first and last dose of open-label study drug plus 1 day and then converted to weeks. There are 5 expected doses during the open-label portion of the study.

A data listing will be provided for treatment adherence during the RDBPC portion and the open-label portion for the safety population.

## 7.  ENDPOINT EVALUATION

### 7.1.  Overview of Efficacy Analysis Methods

Hypothesis testing for all efficacy endpoints will be performed in the mITT population unless otherwise specified. For each estimate that is provided, a corresponding 95% confidence interval will be included.

#### 7.1.1.  Multicenter Studies

Participants were enrolled at 8 sites. Sub-group analyses of the primary endpoint will be performed using each site as a sub-group (see Section 7.2.3).

#### 7.1.2.  Assessment Time Windows

Visit windows for all scheduled visits are provided in Protocol Section 8 and in the Schedule of Events (Protocol Appendix A). A visit that occurred outside of the specified visit window will be recorded as a non-major protocol deviation. Any data collected outside of a visit window will still be included in any analyses of the scheduled visit.

Data collected during unscheduled visits, outside the visit window, or within the appropriate visit window but not on the date of study drug administration (as applicable) will be included in listings.

### 7.2.  Primary Endpoint

#### 7.2.1.  Computation of the Primary Endpoint

*S. aureus* abundance as measured by microbial DNA (*femA* qPCR) on lesional skin at Day 28.

### 7.2.2.    Primary Analysis of the Primary Endpoint

The primary analysis will compare *S. aureus* abundance between treatment arms at Day 28 using an ANCOVA model applied to the mITT population. Prior to analysis, a log-10 transformation will be applied to *S. aureus* abundance. If any participant to be included in the analysis has a value of zero, the participant's value will be adjusted by adding a value of 1 prior to transformation. To account for the study design induced by the randomization scheme, the model will also include fixed effects for clinical site and disease severity at Day 0, as measured by EASI ≥ 21.1 or < 21.1. To increase statistical efficiency, the model will also adjust for *S. aureus* abundance at Day 0. Such a model will allow an estimation of the adjusted geometric mean ratio of the abundance of *S. aureus* at Day 28 on lesional skin (and corresponding 95% confidence interval) between the dupilumab and placebo arms. A two-sided significance level of 0.05 will be used for this analysis.

### 7.2.3.    Sensitivity Analyses of the Primary Endpoint

Analyses for the primary endpoint will be supported by several analyses:

1. To evaluate the distributional assumptions made in the primary analysis model, a permutation test will be used as a supplementary analysis of the primary analysis (Manly, 1997).
2. Rather than exclude participants who were randomized but do not have *S. aureus* abundance measured at Day 0 and Day 28 (as defined by the mITT population), a multiple imputation technique will be employed to impute the missing primary endpoint of *S. aureus* abundance at Day 0 or Day 28 in the ITT population. The model specified in Section 7.2.2 will be run on each imputed dataset, and results will then be pooled to estimate one geometric mean ratio and corresponding 95% CI.
3. The model specified in Section 7.2.2 will be adjusted to include clinical site as a random effect, rather than a fixed effect.
4. To test if the effect of dupilumab on *S. aureus* abundance at Day 28 is modified by particular sub-groups, separate models similar to the model specified in Section 7.2.2 will be developed to test the two-way interaction between treatment arm and variables defining the sub-groups of interest. Sub-groups that will be investigated include:
   o Clinical site
   o Disease severity at Day 0, as measured by EASI ≥ 21.1 or < 21.1
   o Disease severity at Day 0, as measured above and below the median value of thymus and activation-regulated chemokine (TARC)
5. The model specified in Section 7.3.1 which excludes Days 77 and 112 will be used to test the difference between treatment arms at Day 28.

## 7.3.    Secondary Endpoints

### 7.3.1.    Secondary Endpoint 1

**Variable:** *S. aureus* abundance as measured by microbial DNA (femA qPCR) on lesional skin at Days 0, 3, 7, 14, 21, 42, 77 and 112

**Analysis:** *S. aureus* abundance on lesional skin measured at each RDBPC time point (Days 3, 7, 14, 21, 28, and 42) will be log10-transformed and analyzed using a linear (or nonlinear, if appropriate) mixed  model for repeated measures that will include fixed effects for *S. aureus* abundance on lesional skin at Day 0, treatment arm, clinical site, disease severity at Day 0 as measured by EASI ≥ 21.1 or < 21.1, time point of measurement (as a categorical variable), and an interaction term

between treatment arm and time point of measurement. If any participant has a value of zero, the participant's value will be adjusted by adding a value of 1 prior to transformation. The model will be fit using restricted maximum likelihood, and each fixed effect will be tested using an F-test statistic, the Kenward-Roger approximation for degrees of freedom, an unstructured covariance matrix, and a two-sided significance level of 0.05. The model will be used to estimate the following at Days 3, 7, 14, 21, and 42:

- The geometric mean of *S. aureus* abundance on lesional skin in each treatment arm.
- The geometric mean ratio of *S. aureus* abundance on lesional skin between treatment arms.

Additionally, a linear (or nonlinear, if appropriate) mixed model similar to the model specified above will be fit for every time point measured (including the open-label portion of the study). The model will be used to estimate the following at Days 77 and 112:

- The geometric mean of *S. aureus* abundance on lesional skin in each treatment arm
- The geometric mean ratio of *S. aureus* abundance on lesional skin between treatment arms
- The geometric mean ratio of *S. aureus* abundance on lesional skin between Day 77 and Day 42 and between Day 112 and Day 42 in each treatment arm

### 7.3.2.    Secondary Endpoint 2

**Variable:** *S. aureus* abundance as measured by microbial DNA (femA qPCR) on non-lesional skin at Days 0, 3, 7, 14, 21, 28, 42, 77 and 112

**Analysis:** Using participants in the mITT population, *S. aureus* abundance on non-lesional skin will be log10-transformed and will be evaluated by applying two linear (or nonlinear, if appropriate) mixed models for repeated measures (similar to Section 7.3.1). If any participant to be included in the analysis has a value of zero, the participant's value will be adjusted by adding a value of 1 prior to transformation. The model excluding Days 77 and 112 will be used to estimate the following at Days 3, 7, 14, 21, 28, and 42:

- The geometric mean of *S. aureus* abundance on non-lesional skin in each treatment arm.
- The geometric mean ratio of *S. aureus* abundance on non-lesional skin between treatment arms.

The model including Days 77 and 112 will also be used to estimate the following at Days 77 and 112:

- The geometric mean of *S. aureus* abundance on non-lesional skin in each treatment arm
- The geometric mean ratio of *S. aureus* abundance on non-lesional skin between treatment arms
- The geometric mean ratio of *S. aureus* abundance on lesional skin between Day 77 and Day 42 and between Day 112 and Day 42 in each treatment arm

### 7.3.3.    Secondary Endpoint 3

**Variable:** Basal (prior to tape stripping) transepidermal water loss (TEWL) of non-lesional and lesional skin at Days 0, 3, 7, 14, 21, 28, 42, 77 and 112

**Analysis:** Basal TEWL of non-lesional skin will be evaluated by applying two linear (or nonlinear, if appropriate) mixed models for repeated measures (similar to Section 7.3.1) that will include random effects for participant and fixed effects for treatment arm, basal TEWL at Day 0, clinical site, and disease severity at Day 0 as measured by EASI $\geq 21.1$ or $< 21.1$. The model excluding Days 77 and 112 will be used to estimate the following at Days 3, 7, 14, 21, 28, and 42:

- The mean basal TEWL of non-lesional skin in each treatment arm
- The mean difference between treatment arms in basal TEWL of non-lesional skin

The model including Days 77 and 112 will also be used to estimate the following at Days 77 and 112:

- The mean basal TEWL of non-lesional skin in each treatment arm
- The mean difference in basal TEWL of non-lesional skin between treatment arms
- The mean difference in basal TEWL of non-lesional skin between Day 77 and Day 42 and between Day 112 and Day 42 in each treatment arm

The same analyses will be performed using basal TEWL of lesional skin.

### 7.3.4.    Secondary Endpoint 4

**Variable:** TEWL area under the curve (AUC) of non-lesional skin at Days 0, 7, 14, 21, 28, 42, 77 and 112. TEWL is assessed prior to tape stripping and repeated after 5, 10, and 15 tape strips. TEWL AUC will be computed for each participant's non-lesional skin assessment using the trapezoidal rule over the course of their tape strips.

**Analysis:** Using participants in the mITT population, TEWL AUC of non-lesional skin will be evaluated by applying two linear (or nonlinear, if appropriate) mixed models for repeated measures (similar to Section 7.3.1) that will include random effects for participant and fixed effects for treatment arm, TEWL AUC at Day 0, clinical site, and disease severity at Day 0 as measured by EASI $\geq 21.1$ or $< 21.1$. The model excluding Days 77 and 112 will be used to estimate the following at Days 3, 7, 14, 21, 28, and 42:

- The mean TEWL AUC in each treatment arm
- The mean difference between treatment arms in TEWL AUC

The model including Days 77 and 112 will also be used to estimate the following at Days 77 and 112:

- The mean TEWL AUC in each treatment arm
- The mean difference in TEWL AUC between treatment arms
- The mean difference in TEWL AUC between Day 77 and Day 42 and between Day 112 and Day 42 in each treatment arm

### 7.3.5. Secondary Endpoint 5

**Variable:** Change in TEWL per every 5 tape strips (i.e. slope) on non-lesional skin at Days 0, 7, 14, 21, 28, 42, 77 and 112

**Analysis:** TEWL on non-lesional skin will be evaluated by applying a separate linear mixed model for repeated measures for each day of measurement and will include fixed effects for basal TEWL prior to tape stripping, treatment arm, clinical site, disease severity at Day 0 as measured by EASI ≥ 21.1 or < 21.1, tape strip number (as a categorical variable), and an interaction term between treatment arm and tape strip number. The model will be used to estimate the slope (the interaction term of strip and treatment assignment) in each treatment arm.

### 7.3.6. Secondary Endpoints 6-7

**Variable:** Eczema Area and Severity Index [EASI], Investigator Global Assessment [IGA], SCORing Atopic Dermatitis [SCORAD], and Pruritus numerical rating scale [NRS] at Days 0, 3, 7, 14, 21, 28, 42, 77 and 112.

**Analysis:** EASI (as well as percent change over time in EASI and EASI50), IGA, SCORAD (as well as percent change over time in SCORAD), and Pruritus NRS will each be evaluated by applying two linear (or logistic for EASI50) mixed models for repeated measures (similar to Section 7.3.1). Percent change in EASI and SCORAD will be calculated as $100*(X_{post\text{-}baseline} - X_{Day\ 0}) / X_{Day\ 0}$. EASI50 will be calculated as 1 for participants who have ≥ 50 percent decrease in EASI and 0 for participants who have < 50 percent decrease in EASI. EASI ≥ 21.1 or < 21.1 will not be included as a fixed effect for the endpoints corresponding to EASI. The model excluding Day 72 and 112 will be used to estimate the following:

- The mean EASI, IGA, SCORAD, and Pruritus NRS in each treatment arm
- The mean percent change from baseline in EASI and SCORAD in each treatment arm
- The mean difference in EASI, IGA, SCORAD, and Pruritus NRS between treatment arms
- The mean difference in percent change from baseline in EASI and SCORAD between treatment arms
- The odds ratio comparing ≥ 50 percent change in EASI between the treatment arms

The model including Days 77 and 112 will also be used to estimate the following at Days 77 and 112:

- The mean EASI, IGA, SCORAD, and Pruritus NRS in each treatment arm
- The mean difference in EASI, IGA, SCORAD, and Pruritus NRS between treatment arms
- The mean difference in EASI, IGA, SCORAD, and Pruritus NRS between Day 77 and Day 42 and between Day 112 and Day 42 in each treatment arm
- The mean percent change from baseline in EASI  and SCORAD in each treatment arm
- The mean difference in percent change from baseline in EASI and SCORAD between treatment arms
- The mean difference in percent change between Day 77 and Day 42 and between Day 112 and Day 42 in EASI and SCORAD in each treatment arm
- The odds ratio comparing ≥ 50 percent change in EASI between the treatment arms

- The odds ratio comparing ≥ 50 percent change in EASI between Day 77 and Day 42 and between Day 112 and Day 42 in each treatment arm

### 7.3.7. Supportive Analyses of Secondary Endpoints

To compare the cross-sectional and longitudinal effects of *S. aureus* abundance on disease severity between treatment arms, a generalized linear mixed model similar to that described in Chapter 15.4 of Fitzmaurice et al (2004) will be used. By including an interaction between treatment arm and *S. aureus* abundance at Day 0, as well as an interaction between treatment arm and the change in *S. aureus* abundance over time, this model will allow simultaneous estimation of the effect of *S. aureus* abundance at Day 0 on disease severity at Day 0 in the dupilumab and placebo arms separately, as well as the change in *S. aureus* abundance over time on the change in disease severity over time in the dupilumab and placebo arms, separately. The model will also be adjusted for clinical site, skin collection site, and disease severity at Day 0, as measured by EASI ≥ 21.1 or < 21.1 (for endpoints other than EASI).

## 7.4. Exploratory Endpoints

### 7.4.1. Exploratory Endpoint 1

**Variable:** Composition of bacterial taxa in lesional and non-lesional skin at Days 0, 3, 7, 14, 21, 28, 42, 77 and 112.

**Analysis:** See Protocol Section 13.4.5.1.1 for analysis plan.

### 7.4.2. Exploratory Endpoint 2

**Variable:** Abundance of bacterial taxa in lesional and non-lesional skin at Days 0, 3, 7, 14, 21, 28, 42, 77 and 112.

**Analysis:** See Protocol Section 13.4.5.1.1 for analysis plan.

### 7.4.3. Exploratory Endpoint 3

**Variable:** Gene expression in the skin transcriptome in non-lesional skin at Days 0 and 7, and in lesional skin at Days 0, 7, and 21.

**Analysis:** See Protocol Section 13.4.5.2 for analysis plan.

### 7.4.4. Exploratory Endpoint 4

**Variable:** Lipid profiles of non-lesional skin at Days 0, 7, 14, 21, 28, 42, 77 and 112 and of lesional skin at Days 0, 14, 28, and 112

**Analysis:** For each time point, we will compare lipid profiles between dupilumab and placebo arms as previously described (Janssens et al, 2012). Two linear mixed models (one for each collection site), similar to that described in Section 7.3.1, will be used to investigate the effect of dupilumab on lipid profiles over time.

### 7.4.5.    Exploratory Endpoint 5

**Variable:** Expression of *S. aureus* superantigens (SAg), toxins, lipase, and proteases on lesional and non-lesional skin at Days 0, 3, 7, 14, 21, 28, 42, 77 and 112

**Analysis:** We will compare the expression of each *S. aureus* SAg, toxin, lipase, and protease over time between dupilumab and placebo arms using two linear mixed models for repeated measures for each collection site, similar to that described in Section 7.3.1.

Additionally, each endpoint will be considered as a binary indicator for presence/absence as the outcome. Two logistic mixed models that will include fixed effects for presence/absence of the outcome at Day 0, treatment arm, clinical site, disease severity at Day 0 as measured by EASI ≥ 21.1 or < 21.1, time point of measurement (as a categorical variable), and an interaction term between treatment arm and time point of measurement will be fit for all time points prior to Day 77 and every time point measured (including the open-label portion of the study). The models will be used to estimate the odds ratio of presence of the outcome between treatment arms at each time point.

### 7.4.6.    Exploratory Endpoint 6

**Variable:** Confocal imaging of tight junctions (TJs) and relationship to LCs in the epidermis from non-lesional skin at Days 0, 7, and 21.

**Analysis:** The number of LCs in the epidermis in the dupilumab and placebo treatment arms will be compared at Days 0, 7, and 21 using a two-sample t-test. If the variances are equal a two-sample t-test assuming equal variances will be used. If the variances are not equal a two sample t-test assuming unequal variances will be used. The mean number of LC's in the epidermis in each treatment arm and the mean difference between treatment arms will be estimated.

### 7.4.7.    Exploratory Endpoint 7

**Variable:** Percent of coagulase-negative staphylococci [CoNS] isolates that kill *S. aureus* on lesional and non-lesional skin at Days 0, 3, 7, 14, 21, 28, 42, 77 and 112

**Analysis:** The frequency of commensal Staphylococcus species producing antimicrobial activity, as measured by the number of colonies with antimicrobial activity against *S. aureus* per sample, between dupilumab and placebo will be compared using two linear mixed models for repeated measures for each collection site, similar to that described in Sections 7.3.1.

### 7.4.8.    Exploratory Endpoint 8

**Variable:** Peripheral Blood Mononuclear Cell (PBMC) immunoprofiling at Days 0, 14 and 28.

**Analysis:** See Protocol Section 13.4.5.8 for analysis plan.

### 7.4.9.    Exploratory Endpoint 9

**Variable:** Levels of serum biomarkers (e.g. T-helper 2 (Th2) biomarkers) at Days 0, 7, 14, 21, 28, 42, 77 and 112.

**Analysis:** We will compare serum biomarkers over time between dupilumab and placebo arms using two linear mixed models for repeated measures, similar to that described in Section 7.3.1.

### 7.4.10. Exploratory Endpoint 10

**Variable:** The presence of single nucleotide polymorphisms (SNPs).

**Analysis:** See Protocol Section 13.4.5.8 for analysis plan.

## 8. SAFETY EVALUATION

### 8.1. Overview of Safety Analysis Methods

The safety analyses will be performed in the safety population defined in Section 4. Safety assessment summaries will include:

- AEs, defined in Protocol Section 12.2.1.
- Suspected Adverse Reactions (SARs), defined in Protocol Section 12.2.1.1.
- AEs leading to discontinuation of study drug
- AEs leading to withdrawal
- Serious adverse events (SAEs), defined in Protocol Section 12.2.3.
- Deaths
- Clinical laboratory results
- Vital signs
- Physical examinations

These analyses will not be stratified by clinical site.

A TEAE will be defined as any AE not present prior to the initiation of study drug or any AE already present that worsens in either intensity or frequency following exposure to study drug. Any change in clinical status, electrocardiograms (ECGs), routine labs, x-rays, physical examinations, etc. during treatment periods will be considered a treatment-emergent AE. TEAEs will be summarized in the safety population.

Partial dates will be imputed for the purposes of defining TEAEs as follows:

- For a missing start day where the month and year are present, the start day will be set to the first day of the month, unless: 1) the first day of the month is before the date of administration of study drug and the month and year are the same as the month and year of the date of administration of study drug; and 2) the end date is on or after the date of administration of study drug or the end date is completely missing, in which case the start day will be set to the day of administration of study drug.
- For a missing start day and month where the year is present, the start day and month will be set to January 1st, unless: 1) January 1st is before the date of administration of study drug and the year is the same as the year of the date of administration of study drug; and 2) the end date is on or after the date of administration of study drug or the end date is completely

missing, in which case the start day and month will be set to that of the date of administration of study drug.

- For a missing end day where the month and year are present, the end day will be set to the last day of the month, unless the month and year are the same as the month and year of the last contact date for the participant, in which case the end day will be set to that of the participant's last contact date.
- For a missing end day and month where the year is present, the end day and month will be set to the participant's last contact date, unless the year of the participant's last contact date is greater than the end year, in which case the end day and month will be set to December 31st.

For grading an abnormal value or result of a clinical or laboratory evaluation (including, but not limited to, a radiograph, an ultrasound, an ECG etc.), a TEAE is defined as an increase in grade from baseline or from the last post-baseline value that doesn't meet grading criteria. Changes in grade from screening to treatment initiation at Day 0 will also be recorded as AEs, but are not treatment-emergent.

Unless otherwise specified, a data listing for all measurements in Section 8 will be provided for the safety population and sorted in order of treatment arm, participant ID, and time of assessment (e.g., visit, time, and/or event).

## 8.2.    Adverse Events

All AEs will be classified by system organ class (SOC) and preferred term, according to a standardized thesaurus (Medical Dictionary for Regulatory Activities [MedDRA]). MedDRA V21.0 or later will be used to classify all AEs.

The severity of AEs will be classified, as applicable, using the criteria set forth in the National Cancer Institute-Common Terminology Criteria for Adverse Events (NCI-CTCAE) Version 4.03.

Each AE will be entered on the eCRF once at the highest severity.

The relationship, or attribution, of an AE to the study therapy regimen or study procedures is defined in Protocol Section 12.3.2.

Overall summary tables will be developed using the safety population to report the number of events and the number and percentage of participants having at least one event in the following categories:

- AEs
- AEs by maximum grade
- AEs by relationship to study drug
- AEs by relationship to study procedure

AEs classified by treatment arm, MedDRA SOC, and preferred term will be summarized in the Safety Population for each of the following categories:

- AEs
- AEs by relationship to study drug

Summary tables will present the total number of events as well as the number and percentage of participants experiencing the events. If a participant experiences the same AE on multiple occasions, the event will be counted once for each occurrence when reporting the number of AEs. When reporting the number of participants experiencing the events, a participant will only be counted once if they experience an event within the particular SOC or preferred term. Percentages will be based on the number of participants in the safety population.

Similar summaries will be generated for SARs (defined in Protocol Section 12.2.1.1), AEs leading to discontinuation of study drug, and AEs leading to study withdrawal separately.

Data listings will be provided separately for AEs, SARs, AEs leading to discontinuation of study drug, and cases of COVID-19.

## 8.3.     Deaths and Serious Adverse Events

SAEs will be listed and summarized in the same manner described in Section 8.2. Separate displays listing death, including time to death and cause of death, will also be created.

## 8.4.     Clinical Laboratory Evaluation

Clinical laboratory measurements will be performed at Screening and include a complete blood count (CBC) and a complete metabolic panel.  Results will be converted to standardized units where possible.  For numeric laboratory results, descriptive statistics of laboratory values will be presented for each treatment group and overall.  For categorical laboratory results, the number and percentage of participants reporting each result will be presented for each treatment group and overall.  A data listing of all laboratory measurements including an indication of any abnormal values and clinical significance will be provided.

## 8.5.     Vital Signs, Physical Findings, and Other Observations Related to Safety

### 8.5.1.     Vital Signs

Vital sign measurements (including temperature, heart rate, respiratory rate, systolic blood pressure, and diastolic blood pressure) will be collected at each clinic visit. Height and weight will only be measured at the Screening Visit.

A data listing of all vital sign measurements including an indication of any abnormal values and clinical significance will be provided.

### 8.5.2.     Physical Examinations

A complete physical examination will be performed at the Screening Visit and abbreviated physical exams will be performed at all other clinic visits.

A data listing of all physical examination results including an indication of any abnormal observations and clinical significance will be provided.

## 9.    OTHER ANALYSES

### 9.1.    Use of Medications

Medications will be coded according to the latest version of the World Health Organization (WHO) Drug Dictionary. Medications reported on the case report form (CRF) will be categorized for analysis as prior to or after Treatment Initiation (Day 0) by comparing the medication start and stop dates with the date of Treatment Initiation (Day 0). Prior medications will have both the medication start and stop dates prior to the date of Treatment Initiation. Partial and completely missing dates will be imputed for the purposes of classifying concomitant medications as follows:

- Partial dates will be imputed following the same algorithm as above for TEAEs.
- For a missing start date (i.e., day, month, and year are missing), the start date will be set to the date of Screening unless the stop date is prior to the date of Screening, in which case the start date will be set to the stop date.
- For a missing stop date (i.e., day, month, and year are missing), the medication will be treated as ongoing.

The number and percentage of participants receiving prior or concomitant medications will be presented overall and by medication class. When reporting the number of participants receiving the medication, a participant will only be counted once if they ever received the medication within the medication class. Percentages will be based on the number of participants in the safety population.

## 10.    INTERIM ANALYSES AND DATA MONITORING

No interim efficacy or futility analyses are planned for this study. However, statistical analyses of scientific objectives, to include data through Day 112, are planned to occur as soon as applicable data collection is complete. Data will be cleaned and frozen for all visits through Day 112 in order to begin analyses.

No formal interim analysis of safety data will be conducted. The DAIT/NIAID Medical Monitor shall receive monthly reports from the DAIT Statistical And Clinical Coordinating Center (SACCC) compiling new and accumulating information on AEs, SAEs, and pregnancies recorded by the clinical sites on appropriate eCRFs.

The NIAID Allergy and Asthma Data and Safety Monitoring Board (DSMB) will also review safety data at least once per year and on an ad hoc basis as needed. Data for the planned safety reviews will include, at a minimum, a listing of all reported AEs and SAEs that is created by the DAIT SACCC.

## 11.    CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

Not applicable at this time.

## 12. REFERENCES

Ye C, Beyene J, Browne G, et al. Estimating treatment effects in randomised controlled trials with non-compliance: a simulation study. BMJ Open 2014;4:e005362. doi: 10.1136/bmjopen-2014-005362

Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis. John Wiley & Sons, Inc. Hoboken, New Jersey. 2004. 506 pages.

Manly, BFJ. "Randomization". Bootstrap and Monte Carlo Methods in Biology, 2nd Ed. 1997. London: Chapman & Hall.

Janssens M, van Smeden J, Gooris GS, Bras W, Portale G, Caspers PJ, Vreeken RJ, Hankemeier T, Kezic S, Wolterbeek R, Lavrijsen AP, Bouwstra JA. "Increase in short-chain ceramides correlates with an altered lipid organization and decreased barrier function in atopic eczema patients." J Lipid Res. 2012; 53(12):2755-66.

Twisk J, Bosman L, Hoekstra T, Rijnhart J, Welton M, Heymansm. "Different ways to estimate treatment effects in randomized controlled trials." Contemporary Clinical Trials Communications. 2018.